

## PATENT ABSTRACTS OF JAPAN

(11)Publication number : 08-328585

(43)Date of publication of application : 13.12.1996

(51)Int.Cl. G10L 3/00  
 G10L 3/00  
 G06F 17/27

(21)Application number : 07-133762

(71)Applicant : SONY CORP

(22)Date of filing : 31.05.1995

(72)Inventor : ASANO KOJI

WATARI MASAO

AKAHA MAKOTO

KAGAMI TETSUYA

ISHII KAZUO

TANAKA MIYUKI

KATO YASUHIKO

TSUNODA HIROSHI

OGAWA HIROAKI

(54) METHOD AND DEVICE FOR NATURAL LANGUAGE PROCESSING AND  
 METHOD AND DEVICE FOR VOICE RECOGNITION

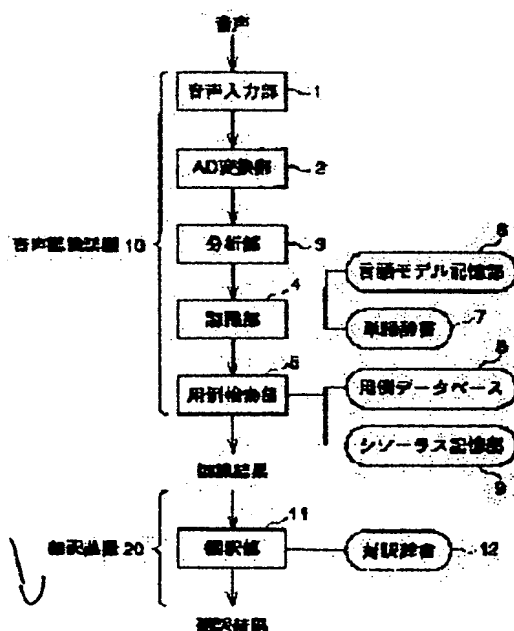
## (57)Abstract:

PURPOSE: To improve the precision in a natural language processing by making possible to select more appropriate utilization examples to process an inputted sentence.

CONSTITUTION: A voice translation device consists of a voice recognition device 10 which recognizes voices and a translation device 20 which translates the voice recognition results into a prescribed language. The device 10 is provided with a utilization example storage means (a utilization example (database) 8) and a selecting means (an example retrieving section) 5. The means 8 stores plural examples. The means 5 computes the degree of similarity between the respective

examples stored in the means 8 and the inputted

sentence and selects the example, which has a highest degree of similarity.



Moreover, the means 5 adds weight based on the context by the selected examples up to now and computes the degree of similarity. Thus, the precision in the degree of similarity is improved, more appropriate examples are selected for the process and the rate of recognition of the voices is improved.

---

## LEGAL STATUS

[Date of request for examination] 14.06.2001

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's decision of rejection]

[Date of extinction of right]

Copyright (C); 1998,2000 Japan Patent Office

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平8-328585

(43) 公開日 平成8年(1996)12月13日

(51) Int.Cl. <sup>8</sup>	識別記号	庁内整理番号	F I	技術表示箇所
G 1 0 L 3/00	5 6 1		G 1 0 L 3/00	5 6 1 G
	5 3 1			5 3 1 G
G 0 6 F 17/27		8420-5L	G 0 6 F 15/38	J

審査請求 未請求 請求項の数10 O L (全 13 頁)

(21) 出願番号 特願平7-133762

(22) 出願日 平成7年(1995)5月31日

(71) 出願人 000002185

ソニー株式会社

東京都品川区北品川6丁目7番35号

(72) 発明者 浅野 康治

東京都品川区北品川6丁目7番35号 ソニー株式会社内

(72) 発明者 渡 雅男

東京都品川区北品川6丁目7番35号 ソニー株式会社内

(72) 発明者 赤羽 誠

東京都品川区北品川6丁目7番35号 ソニー株式会社内

(74) 代理人 弁理士 稲本 義雄

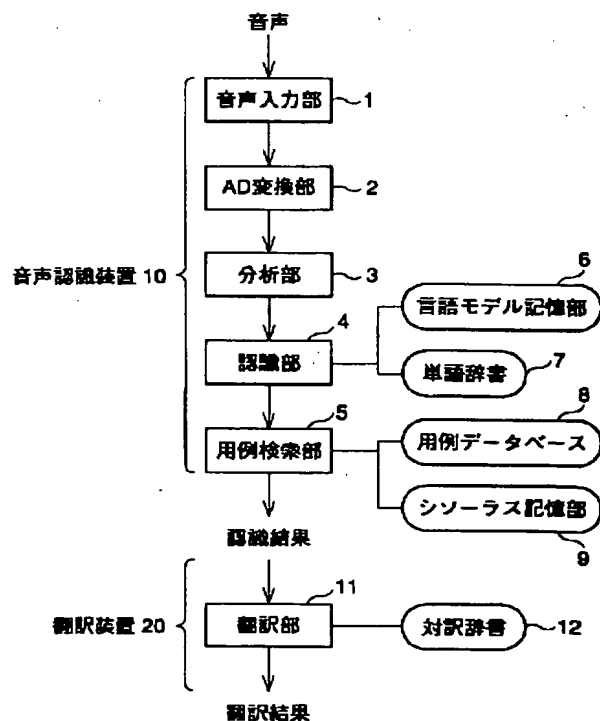
最終頁に続く

(54) 【発明の名称】 自然言語処理装置および自然言語処理方法、並びに音声認識装置および音声認識方法

(57) 【要約】

【目的】 音声の認識率を向上させる。

【構成】 用例検索部5では、用例データベース8に記憶されている複数の用例それぞれと、認識部4から出力される複数の認識結果候補との類似度が計算され、その類似度を最も高くする用例が選択され、その用例を用いて、最終的な音声認識結果が求められる。この場合において、用例検索部5は、いままでに選択した用例による文脈に基づいて重みを付けて、類似度を計算する。



**【特許請求の範囲】**

**【請求項 1】** 入力文を、その入力文と最も類似する用例を用いて処理する自然言語処理装置であって、複数の前記用例を記憶している用例記憶手段と、前記用例記憶手段に記憶されている用例それぞれと、前記入力文との類似度を計算し、その類似度を最も高くする前記用例を選択する選択手段とを備え、前記選択手段は、いままでに選択した前記用例による文脈に基づいて重みを付けて、前記類似度を計算することを特徴とする自然言語処理装置。

**【請求項 2】** 前記用例記憶手段は、次の処理に用いられる可能性のある前記用例を予測するための予測情報をさらに記憶しており、前記選択手段は、前記予測情報に基づいて、前記重みを付けることを特徴とする請求項 1 に記載の自然言語処理装置。

**【請求項 3】** 前記用例記憶手段は、前記予測情報として、前記用例を次の処理に用いられる可能性のある前記用例と対応付けたものを記憶しており、前記選択手段は、前回の処理で用いられた前記用例と対応付けられている前記用例に対する前記類似度が高くなるように、または前回の処理で用いられた前記用例と対応付けられていない前記用例に対する前記類似度が低くなるように、前記重みを付けることを特徴とする請求項 2 に記載の自然言語処理装置。

**【請求項 4】** 前記用例記憶手段は、次の処理に用いられる可能性のある前記用例を学習により求めたものを、前記用例と対応付けて記憶していることを特徴とする請求項 3 に記載の自然言語処理装置。

**【請求項 5】** 前記用例記憶手段は、前記予測情報を、所定の場面ごとに用いるべきものに分類して記憶しており、前記選択手段は、場面に応じた前記予測情報を用いることを特徴とする請求項 2 に記載の自然言語処理装置。

**【請求項 6】** 複数の用例を記憶している用例記憶手段を備え、入力文を、その入力文と最も類似する前記用例を用いて処理する自然言語処理装置の自然言語処理方法であって、前記用例記憶手段に記憶されている用例それぞれと、前記入力文との類似度を計算し、その類似度を最も高くする前記用例を選択するときに、いままでに選択した前記用例による文脈に基づいて重みを付けて、前記類似度を計算することを特徴とする自然言語処理方法。

**【請求項 7】** 文音声認識を行う音声認識装置であって、音声音を音響分析し、その特徴パラメータを抽出する抽出手段と、前記抽出手段より出力される前記特徴パラメータに基づいて前記音声の認識を行い、複数の認識結果候補と、それぞれの尤度とを求める認識結果候補算出手段と、

複数の用例を記憶している用例記憶手段と、前記複数の認識結果候補それぞれと、前記用例記憶手段に記憶されている用例それぞれとの言語的な類似度である言語的類似度を計算する言語的類似度計算手段と、前記複数の認識結果候補それぞれと、前記用例記憶手段に記憶されている用例それぞれとの総合的な類似度である総合類似度を、前記言語的類似度および尤度に基づいて計算する総合類似度計算手段と、前記総合類似度を最も高くする前記用例を選択する選択手段と、前記選択手段により選択された前記用例に基づいて、前記音声の認識結果を求める認識結果算出手段とを備えることを特徴とする音声認識装置。

**【請求項 8】** 前記総合類似度計算手段は、前記言語的類似度と尤度とを、所定の割合で加算することにより前記総合類似度を計算することを特徴とする請求項 7 に記載の音声認識装置。

**【請求項 9】** 前記総合類似度計算手段は、前記言語的類似度と尤度とを加算するときの前記所定の割合を動的に変化させることを特徴とする請求項 8 に記載の音声認識装置。

**【請求項 10】** 複数の用例を記憶している用例記憶手段を備え、文音声認識を行う音声認識装置の音声認識方法であって、音声音を音響分析し、その特徴パラメータを抽出し、前記特徴パラメータに基づいて前記音声の認識を行い、複数の認識結果候補と、それぞれの尤度とを求め、前記複数の認識結果候補それぞれと、前記用例記憶手段に記憶されている用例それぞれとの言語的な類似度である言語的類似度を計算し、前記複数の認識結果候補それぞれと、前記用例記憶手段に記憶されている用例それぞれとの総合的な類似度である総合類似度を、前記言語的類似度および尤度に基づいて計算し、前記総合類似度を最も高くする前記用例を選択し、その用例に基づいて、前記音声の認識結果を求めることを特徴とする音声認識方法。

**【発明の詳細な説明】****【0001】**

**【産業上の利用分野】** 本発明は、例えば文字や音声で入力された言語を処理し、その認識や翻訳などを行う場合などに用いて好適な自然言語処理装置および自然言語処理方法、並びに音声認識装置および音声認識方法に関する。

**【0002】**

**【従来の技術】** 従来より、例えばコンピュータなどを用いて人間の用いる文字言語、音声言語の処理（以下、適宜、これらを総称して自然言語処理という）を行うための研究が数多くなされている。しかしながら、これらの研究の多くで採用されている自然言語を処理する方法

は、専門家があらかじめ自然言語において現れる様々な言語現象を文法規則という抽象化した形で記述し、その規則に基づいて処理を行うというものであったため、文法規則の記述の困難性に起因する問題があった。

【0003】即ち、人間どうしの行うコミュニケーションの過程で現れる様々な言語現象を、完全に網羅するような抽象的な文法規則を記述することは、専門家であっても、ほとんど不可能に近く、このため、文法規則によって網羅されていない言語現象が現れた場合には、誤った処理が行われることがあった。そして、このような文法規則によって網羅されていない言語現象が見つかった場合、文法規則をどのように修正すれば良いかは一般に分かりにくく、このため、文法規則の修正によって、これまで正常に処理されていた言語現象が処理できなくなるなどの弊害が生じる問題があった。

【0004】そこで、最近では、このような弊害を根本的に解消すべく、用例を利用した自然言語処理の研究が盛んに行われている。これは、抽象的な文法規則に代えて、大量の用例を登録した用例データベースを用意し、処理すべきテキストデータなどが入力されると、用例データベースの中から、入力されたテキストデータに類似している用例を検索し、その用例に基づいて自然言語処理を行うというものである。

【0005】例えば、特開平 3-276367 号公報には、このような自然言語処理を適用した用例主導型機械翻訳装置が開示されている。この翻訳装置では、原文の用例とその対訳例との組を、用例データベースに多数登録しておき、入力文として、所定の言語で記述された原文が入力されると、その原文に最も類似して用例が用例データベースから検索され、その結果得られた用例の対訳例にしたがって、原文の翻訳が行われるようになされている。

【0006】また、用例を利用した自然言語処理を適用した音声言語処理では、通常の連続音声認識装置を用い、入力された音声の音声認識結果を確定した後、その音声認識結果を入力文として、これに最も類似する用例を、用例データベースから検索し、その用例を用いて、翻訳などが行われるようになされている。

【0007】

【発明が解決しようとする課題】ところで、人間がコミュニケーションの手段として自然言語を用いる場合には、個々の文が有する意味には、文脈あるいは対話履歴というべき、それまでの話の流れが反映されることが普通であり、従って、この文脈（対話履歴）は、自然言語処理にあたって、重要な要素であると考えられる。

【0008】しかしながら、上述したような、従来の用例を利用する方法では、入力文を処理するのに用いる用例を検索する際の、入力文と用例との類似度の計算が、単語を、その意味の類似性（概念）に基づいて木構造に階層化したシソーラスを用いて、単語のシソーラス上で

の意味上の類似性のみを対象にして行われるようになされており、文脈は考慮されていなかった。このため、入力文を処理するのに真に適切な用例が検索されるかどうかが不明であった。

【0009】また、従来の音声言語処理においては、他の自然言語処理の前に音声認識処理を行い、その結果得られる音声認識結果を確定した後、その音声認識結果に最も類似する用例を検索し、その用例を用いて、機械翻訳などの自然言語処理が施されるようになされているため、音声認識結果が誤っている場合には、正しい翻訳結果を得ることが困難となる課題があった。

【0010】さらに、従来の音声言語処理では、音声認識処理の結果得られる音声認識結果候補の、いわば確からしさというべき尤度が考慮されておらず、やはり入力文を処理するのに真に適切な用例が検索されるかどうか不明であった。

【0011】本発明は、このような状況に鑑みてなされたものであり、自然言語処理に際し、入力文を処理するのに、より適切な用例を選択することができるようにし、これにより自然言語処理の精度を向上させることができるようにするものである。

【0012】

【課題を解決するための手段】本発明の自然言語処理装置は、複数の用例を記憶している用例記憶手段と、用例記憶手段に記憶されている用例それぞれと、入力文との類似度を計算し、その類似度を最も高くする用例を選択する選択手段とを備え、選択手段が、いままでに選択した用例による文脈に基づいて重みを付けて、類似度を計算することを特徴とする。

【0013】この自然言語処理装置においては、用例記憶手段に、次の処理に用いられる可能性のある用例を予測するための予測情報をさらに記憶させておき、選択手段には、予測情報に基づいて、重みを付けさせることができる。また、用例記憶手段には、予測情報として、用例を次の処理に用いられる可能性のある用例と対応付けたものを記憶させておき、選択手段には、前回の処理で用いられた用例と対応付けられている用例に対する類似度が高くなるように、または前回の処理で用いられた用例と対応付けられていない用例に対する類似度が低くなるように、重みを付けさせることができる。用例記憶手段には、次の処理に用いられる可能性のある用例を学習により求めたものを、用例と対応付けて記憶させておくことができる。また、用例記憶手段には、予測情報を、所定の場面ごとに用いるべきものに分類して記憶させておき、選択手段には、場面に応じた予測情報を用いさせることができる。

【0014】本発明の自然言語処理方法は、複数の用例を記憶している用例記憶手段を備え、入力文を、その入力文と最も類似する用例を用いて処理する自然言語処理装置の自然言語処理方法であって、用例記憶手段に記憶

されている用例それぞれと、入力文との類似度を計算し、その類似度を最も高くする用例を選択するときに、いままでに選択した用例による文脈に基づいて重みを付けて、類似度を計算することを特徴とする。

【0015】本発明の音声認識装置は、音声を音響分析し、その特徴パラメータを抽出する抽出手段と、抽出手段より出力される特徴パラメータに基づいて音声の認識を行い、複数の認識結果候補と、それぞれの尤度とを求める認識結果候補算出手段と、複数の用例を記憶している用例記憶手段と、複数の認識結果候補それぞれと、用例記憶手段に記憶されている用例それぞれの言語的な類似度である言語的類似度を計算する言語的類似度計算手段と、複数の認識結果候補それぞれと、用例記憶手段に記憶されている用例それぞれの総合的な類似度である総合類似度を、言語的類似度および尤度に基づいて計算する総合類似度計算手段と、総合類似度を最も高くする用例を選択する選択手段と、選択手段により選択された用例に基づいて、音声の認識結果を求める認識結果算出手段とを備えることを特徴とする。

【0016】この音声認識装置においては、総合類似度計算手段には、言語的類似度と尤度とを、所定の割合で加算することにより総合類似度を計算させることができる。また、総合類似度計算手段には、言語的類似度と尤度とを加算するときの所定の割合を動的に変化させることができる。

【0017】本発明の音声認識方法は、複数の用例を記憶している用例記憶手段を備え、文音声認識を行う音声認識装置の音声認識方法であって、音声を音響分析し、その特徴パラメータを抽出し、特徴パラメータに基づいて音声の認識を行い、複数の認識結果候補と、それぞれの尤度とを求め、複数の認識結果候補それぞれと、用例記憶手段に記憶されている用例それぞれの言語的な類似度である言語的類似度を計算し、複数の認識結果候補それぞれと、用例記憶手段に記憶されている用例それぞれの総合的な類似度である総合類似度を、言語的類似度および尤度に基づいて計算し、総合類似度を最も高くする用例を選択し、その用例に基づいて、音声の認識結果を求めることを特徴とする。

【0018】

【作用】本発明の自然言語処理装置においては、用例記憶手段は、複数の用例を記憶しており、選択手段は、用例記憶手段に記憶されている用例それぞれと、入力文との類似度を計算し、その類似度を最も高くする用例を選択するようになされている。さらに、選択手段は、いままでに選択した用例による文脈に基づいて重みを付けて、類似度を計算するようになされている。

【0019】本発明の自然言語処理方法においては、用例記憶手段に記憶されている用例それぞれと、入力文との類似度を計算し、その類似度を最も高くする用例を選択するときに、いままでに選択した用例による文脈に基

づいて重みを付けて、類似度を計算するようになされている。

【0020】本発明の音声認識装置においては、抽出手段は、音声を音響分析し、その特徴パラメータを抽出し、認識結果候補算出手段は、抽出手段より出力される特徴パラメータに基づいて音声の認識を行い、複数の認識結果候補と、それぞれの尤度とを求めるようになされている。用例記憶手段は、複数の用例を記憶しており、言語的類似度計算手段は、複数の認識結果候補それぞれと、用例記憶手段に記憶されている用例それぞれの言語的な類似度である言語的類似度を計算するようになされている。総合類似度計算手段は、複数の認識結果候補それぞれと、用例記憶手段に記憶されている用例それぞれの総合的な類似度である総合類似度を、言語的類似度および尤度に基づいて計算し、選択手段は、総合類似度を最も高くする用例を選択するようになされている。認識結果算出手段は、選択手段により選択された用例に基づいて、音声の認識結果を求めるようになされている。

【0021】本発明の音声認識方法においては、音声を音響分析し、その特徴パラメータを抽出し、特徴パラメータに基づいて音声の認識を行い、複数の認識結果候補と、それぞれの尤度とを求め、複数の認識結果候補それぞれと、用例記憶手段に記憶されている用例それぞれの言語的な類似度である言語的類似度を計算し、複数の認識結果候補それぞれと、用例記憶手段に記憶されている用例それぞれの総合的な類似度である総合類似度を、言語的類似度および尤度に基づいて計算し、総合類似度を最も高くする用例を選択し、その用例に基づいて、音声の認識結果を求めるようになされている。

【0022】

【実施例】以下に、本発明の実施例を説明するが、その前に、特許請求の範囲に記載の発明の各手段と以下の実施例との対応関係を明らかにするために、各手段の後の括弧内に、対応する実施例（但し、一例）を付加して、本発明の特徴を記述すると、次のようになる。

【0023】即ち、請求項1に記載の自然言語処理装置は、入力文を、その入力文と最も類似する用例を用いて処理する自然言語処理装置であって、複数の用例を記憶している用例記憶手段（例えば、図1に示す用例データベース8など）と、用例記憶手段に記憶されている用例それぞれと、入力文との類似度を計算し、その類似度を最も高くする用例を選択する選択手段（例えば、図1に示す用例検索部5など）とを備え、選択手段は、いままでに選択した用例による文脈に基づいて重みを付けて、類似度を計算することを特徴とする。

【0024】請求項6に記載の自然言語処理方法は、複数の用例を記憶している用例記憶手段（例えば、図1に示す用例データベース8など）を備え、入力文を、その入力文と最も類似する用例を用いて処理する自然言語処

理装置の自然言語処理方法であって、用例記憶手段に記憶されている用例それぞれと、入力文との類似度を計算し、その類似度を最も高くする用例を選択するときに、いままでに選択した用例による文脈に基づいて重みを付けて、類似度を計算することを特徴とする。

【0025】請求項7に記載の音声認識装置は、文音声認識を行う音声認識装置であって、音声を音響分析し、その特徴パラメータを抽出する抽出手段（例えば、図5に示す分析部3など）と、抽出手段より出力される特徴パラメータに基づいて音声の認識を行い、複数の認識結果候補と、それぞれの尤度とを求める認識結果候補算出手段（例えば、図5に示す認識部14など）と、複数の用例を記憶している用例記憶手段（例えば、図5に示す用例データベース8など）と、複数の認識結果候補それぞれと、用例記憶手段に記憶されている用例それぞれとの言語的な類似度である言語的類似度を計算する言語的類似度計算手段（例えば、図6に示すプログラムの処理ステップS12など）と、複数の認識結果候補それぞれと、用例記憶手段に記憶されている用例それぞれとの総合的な類似度である総合類似度を、言語的類似度および尤度に基づいて計算する総合類似度計算手段（例えば、図6に示すプログラムの処理ステップS13など）と、総合類似度を最も高くする用例を選択する選択手段（例えば、図6に示すプログラムの処理ステップS14など）と、選択手段により選択された用例に基づいて、音声の認識結果を求める認識結果算出手段（例えば、図5に示す用例検索部15など）とを備えることを特徴とする。

【0026】請求項10に記載の音声認識方法は、複数の用例を記憶している用例記憶手段（例えば、図5に示す用例データベース8など）を備え、文音声認識を行う音声認識装置の音声認識方法であって、音声を音響分析し、その特徴パラメータを抽出し、特徴パラメータに基づいて音声の認識を行い、複数の認識結果候補と、それぞれの尤度とを求め、複数の認識結果候補それぞれと、用例記憶手段に記憶されている用例それぞれとの言語的な類似度である言語的類似度を計算し、複数の認識結果候補それぞれと、用例記憶手段に記憶されている用例それぞれとの総合的な類似度である総合類似度を、言語的類似度および尤度に基づいて計算し、総合類似度を最も高くする用例を選択し、その用例に基づいて、音声の認識結果を求めることを特徴とする。

【0027】なお、勿論この記載は、各手段を上記したものに限定することを意味するものではない。

【0028】図1は、本発明を適用した音声翻訳装置の一実施例の構成を示している。なお、この音声翻訳装置は、大きく分けて、音声を認識する音声認識装置10と、その音声認識結果を、所定の言語に翻訳する翻訳装置20とから構成されている。そして、音声認識装置10は、例えば日本語および英語の両方の音声の文音声認

識（会話音声認識、あるいは音声理解）を行うことができるようになされており、翻訳装置20も、音声認識結果を、日本語および英語の両方に翻訳することができるようになされている。従って、この音声翻訳装置は、日本語で入力された音声の英語への翻訳、および英語で入力された音声の日本語への翻訳の両方を行うことができるようになされている。

【0029】音声入力部1は、そこに入力される、いわば空気の波である音声を、電気的な波である音声信号に変換する、例えばマイクなどと、そのマイクから出力される音声信号を増幅するアンプなどで構成されている。AD変換部2は、音声入力部1から出力されるアナログの音声信号を、所定のクロックのタイミングでサンプリングして量子化を行い、デジタルの音声信号（デジタルデータ）に変換するようになされている。

【0030】分析部3は、AD変換部2より出力される音声信号を音響分析し、これにより、例えば所定の帯域ごとの音声のパワーや、線形予測係数（LPC）、ケプストラム係数などの音声の特徴パラメータを抽出するようになされている。即ち、分析部3は、例えばフィルタバンクにより、音声信号を所定の帯域ごとにフィルタリングし、そのフィルタリング結果を整流平滑化することで、所定の帯域ごとの音声のパワーを求めるようになされている。あるいは、分析部3は、入力された音声に対し、例えば線形予測分析処理を施すことで、線形予測係数を求め、またその線形予測係数からケプストラム係数を求めるようになされている。

【0031】分析部3で求められた特徴パラメータは、そのまま、あるいは、そこで必要に応じてベクトル量子化されて、認識部4に出力されるようになされている。

【0032】認識部4は、分析部3からの特徴パラメータ（あるいは、特徴パラメータをベクトル量子化して得られるシンボル）に基づき、例えばダイナミックプログラミング（DP）マッチング法や、隠れマルコフモデル（HMM）などの音声認識アルゴリズムにしたがい、後述する言語モデル記憶部6および単語辞書7を参照して、例えば文音声認識を行い、複数の文単位の音声認識結果候補（以下、適宜、文候補という）を求めて、用例検索部5に出力するようになされている。

【0033】用例検索部5は、認識部4から供給される複数の文候補それぞれを、例えばそこから助詞を削除することなどにより、自立語のみでなる単語列に変換するようになされている。そして、用例検索部5は、その複数の文候補それぞれの単語列（以下、適宜、入力単語列という）と、後述する用例データベース8に記憶されている用例それぞれとの類似度を計算し、その類似度を最も高くする用例を選択するようになされている。そして、その選択した用例に基づいて、音声入力部1に入力された音声の認識結果（文）を決定するようになされている。

【0034】なお、用例検索部5における類似度の計算は、後述するようにシソーラス記憶部9に記憶されているシソーラスを用い、いままでに選択した用例による文脈に基づいて重みを付けて行われるようになされている。

【0035】言語モデル記憶部6は、例えばバイグラム(Bigram)やトリグラム(Trigram)などの統計的言語モデルなどを記憶しており、認識部4は、上述したような音声認識処理を、言語モデル記憶部6に記憶されている言語モデルにより緩い言語的制約をかけ、その制約の下、例えばビタビ(Viterbi)アルゴリズムを用いたビームサーチなどにより、適当に枝刈しながら行うようになされている。そして、その結果残った複数の音声認識結果候補としての複数の文候補を、用例検索部5に出力するようになされている。

【0036】単語辞書7には、音声認識の対象とする単語の見出し(例えば、「発明」という単語であれば「発明」という見出し)およびその音韻情報(読み)、さらに必要ならば単語の品詞その他の情報が対応付けられて記憶(登録)されている。なお、単語辞書7には、少なくとも、後述する用例データベース8に記憶されている用例を構成する単語が記憶されている。認識部4では、この単語辞書6に記憶されている単語を対象として、音声認識が行われるようになされている。

【0037】用例データベース8には、複数の用例が記憶(登録)されている。この用例は、例えば新聞に掲載されている文章や、アナウンサーが読み上げる原稿に記載の文章などなどに基づいて作成される。シソーラス記憶部9は、少なくとも単語辞書7に登録されている単語を、その概念ごとに分類して記憶している。本実施例では、シソーラス記憶部9には、後述する図2に示すように、単語を、その概念に基づいて木構造に階層化したシソーラスが記憶されている。

【0038】以上の各ブロックから、音声認識装置10が構成されている。

【0039】一方、翻訳装置20は、翻訳部11および対訳辞書12から構成されている。翻訳部11は、音声認識装置10(用例検索部5)より出力される音声認識結果の言語を、対訳辞書43を参照しながら、他の言語に翻訳するようになされている。なお、翻訳部11は、例えば音声合成装置(図示せず)を内蔵しており、そこで、翻訳結果に対応する合成音を生成して出力するようになされている。対訳辞書43は、例えば日本語の単語と、それを英語に訳した英単語(あるいは英単語列)とを対応付けて記憶している。

【0040】なお、言語モデル記憶部6、単語辞書7、用例データベース8、およびシソーラス記憶部9には、例えば日本語および英語の両方に関するものが記憶されており、これにより、音声認識装置10では、上述したように日本語および英語のいずれの音声も認識すること

ができるようになされている。また、翻訳部11では、日本語または英語の音声認識結果が入力された場合には、それぞれを、英語または日本語に翻訳するようになされている。

【0041】以上のように構成される音声翻訳装置においては、音声入力部1に入力された音声は、AD変換部2、さらには分析部3を介することにより特徴パラメータ(あるいはシンボル)とされて、認識部4に出力される。認識部4では、分析部3の出力を用いて、例えばHMMにしたがい、音声認識処理が行われる。

【0042】ここで、HMMについて簡単に説明する。HMMは、非決定有限状態オートマトンとして定義され、そのモデルは、幾つかの状態と、その状態間の遷移を表すパスから構成される。このようなモデルにおいて、各状態からの状態の遷移過程はマルコフ過程とされ、また、状態が遷移するときにはシンボルが1つ出力されるものとして、モデルの学習が行われる。いま、モデルが有する状態がN個あり、モデル(状態)から出力されるシンボルの種類がK個あるとすると、この学習では、多数の学習データを用い、状態が、状態*i*から状態*j*に遷移する確率(状態遷移確率)  $a_{ij}$ と、そのときにシンボル  $y_k$  が出力される確率(出力シンボル確率)  $b_i(j, y_k)$  が求められる(但し、 $0 < i, j < N+1$ ,  $0 < k < K+1$ )。

【0043】なお、HMMのパラメータには、最初に状態*i*にいる確率(初期状態確率)  $\pi_i$ もあるが、音声認識では、状態が、自分自身か、あるいは自身より右側の状態にしか遷移しないleft-to-rightモデルが、通常用いられるので、初期状態は、モデルの最も左側の状態とされる(最初に、最も左側の状態にいる確率が1とされ、他の状態にいる確率は0とされる)。このため、通常は、学習において、初期状態確率を求める必要はない。

【0044】一方、認識時には、学習の結果得られた状態遷移確率および出力シンボル確率を用いて、分析部3から出力されるシンボル系列が観測(生起)される確率(生起確率)が計算され、その確率の高いものが認識結果候補(あるいは認識結果)とされる。

【0045】本実施例では、認識部4には、あらかじめ学習を行うことにより得られた、例えば音素単位のモデル(音素モデル)が記憶されており、認識部4は、単語辞書7に登録されている単語の音韻情報を参照して、音素モデルを連結し、単語辞書7に登録されている単語のモデルを作成する。そして、このモデルを用いて、上述したように生起確率を計算し、その確率の高い単語を求める。

【0046】なお、認識部4には、音素モデルではなく、例えば単語単位のモデル(単語モデル)を記憶させておき、そのモデルをそのまま用いて、連続音声認識させるようにすることも可能である。



【0047】さらに、認識部4は、上述したような処理を、言語モデル記憶部6に記憶されているバイグラム、トリグラムにより緩い言語的制約をかけ、その制約の下、例えばビタビアルゴリズムを用いたビームサーチなどにより、適当に枝刈しながら行う。この場合、認識部4において、文候補を得るまでのサーチスペースが絞り込まれるので、音声認識装置10の音声認識処理における演算量の低減、さらにはその処理の高速化を図ることができる。

【0048】また、認識部4でかけられる言語モデルによる制約は緩いため、そこで得られる文候補の中には、文法的、意味的に正しいものだけでなく、誤ったものも含まれると考えられるが、そのようなものは、後段の用例検索部5で計算される類似度が低くなるため、最終的な音声認識結果とはされる確率は非常に少ない。

【0049】なお、バイグラム、トリグラムは、例えば1次、2次のマルコフ過程のモデルで、音素、音節、単語などの連鎖確率を大量のテキストデータベースを基にして学習したものであり、自然言語の局所的な性質を精度良く近似することのできるモデルとして知られている。

【0050】また、言語的制約は、言語モデルによる他、例えば有限状態ネットワークなどを用いてかけるようにすることも可能である。

【0051】以上のような言語的制約の下、得られた複数の文候補は、用例検索部5に出力され、上述したように単語列に変換される。さらに、用例検索部5は、複数の単語列（入力単語列）それぞれと、用例データベース8に記憶されている用例それぞれとの類似度を計算し、その類似度を最も高くする用例を選択する。

【0052】ここで、本実施例では、類似度の計算は、例えば、前述した特開平3-276367号に開示されているような、単語を、その意味の類似性（概念）に基づいて木構造に階層化したシソーラスを用いて行われる。即ち、入力単語列を構成する、ある単語と、その単語に対応する、用例を構成する単語とが、同一のカテゴリに属すると考えられる概念の階層が第k階層であった場合に、これらの単語間の概念的な類似性を表す単語類似度を、 $(k-1)/n$ （但し、nは、シソーラスの階層数）とし、入力単語列を構成する単語それぞれと、用例を構成する単語それぞれについての単語類似度を積算する。そして、その積算結果を、入力単語列と、用例との類似度とする。

【0053】具体的には、例えば、いま、シソーラス記憶部9に、図2に示すようなシソーラスが記憶されているものとした場合、以下のようにして類似度が計算される。

【0054】但し、図2においては、長方形で囲んであるものは概念を表し、楕円で囲んであるものが単語を表す。同図では、最も上の階層（第4階層）に属する概念

が、「性状」、「変動」、「人物」、「社会」、「物品」、その他に分類されており、そのうちの、例えば概念「変動」は、それに含まれる概念「増減」、「経過」、その他に分類されている。さらに、例えば概念「経過」は、それに含まれる概念「経過」、「到来」、その他に分類されており、そのうちの、例えば概念「到来」には、その範疇にある単語「訪問する」、その他が属するものとされている。

【0055】また、図2においては、最も下の概念の階層を第1階層とし、下から2番目の概念の階層を第2階層とし、以下同様にして、下から3番目の概念の階層、または最も上の概念の階層を、それぞれ第3階層、または第4階層としている。図2のシソーラスは4階層（の概念）で構成されるから、シソーラスを第1階層までさかのぼることにより概念が一致する単語どうしの単語類似度は $0 = (1-1)/4$ となり、また、シソーラスを第2階層までさかのぼることにより概念が一致する単語どうしの類似度は $1/4 = (2-1)/4$ となる。以下同様に、シソーラスを第3または第4階層までさかのぼることにより概念が一致する単語どうしの単語類似度は $1/2$ または $3/4$ となる。

【0056】例えば、いま、認識部4から、文候補「私は学校へ行く」が用例検索部5に出力され、そこで、この文候補が、上述したように助詞で分離されることにより、入力単語列（「私」、「学校」、「行く」）とされた場合、この入力単語列（「私」、「学校」、「行く」）と、例えば用例「彼は会社を訪問する」との類似度は、次のように計算される。

【0057】即ち、まず入力単語列（「私」、「学校」、「行く」）を構成する単語「私」、「学校」、「行く」それぞれと、用例「彼は会社を訪問する」を構成する、「私」、「学校」、「行く」に対応する単語「彼」、「会社」、「訪問する」それぞれとの単語類似度が計算される。

【0058】ここで、単語XとYとの単語類似度を、 $d(X, Y)$ と表すと、単語「私」と「彼」とは、第2階層までさかのぼることにより概念「人称」に一致するので、単語類似度 $d$ （「私」、「彼」）は $1/4$ となる。また、単語「学校」と「会社」とは、第2階層までさかのぼることにより概念「施設」に一致するので、単語類似度 $d$ （「学校」、「会社」）は $1/4$ となる。さらに、単語「行く」と「訪問する」とは、やはり第2階層までさかのぼることにより概念「経過」に一致するので、単語類似度 $d$ （「行く」、「訪問する」）は $1/4$ となる。

【0059】以上の単語類似度を積算すると、その積算値は $3/4 = (1/4 + 1/4 + 1/4)$ となり、これが、入力単語列（「私」、「学校」、「行く」）と用例「彼は会社を訪問する」との類似度とされる。

【0060】また、この入力単語列（「私」、「学

校」、「行く」と、例えば用例「これは木でできている」との類似度は、次のように計算される。

【0061】即ち、入力単語列（「私」、「学校」、「行く」）を構成する単語「私」、「学校」、「行く」それぞれと、用例「これは木でできている」を構成する、「私」、「学校」、「行く」に対応する単語「これ」、「木」、「できる」（「できている」は「できる」とされる）それぞれとの単語類似度  $d$ （「私」、「これ」）、 $d$ （「学校」、「木」）、 $d$ （「行く」、「できる」）は、上述したようにして、 $3/4$ 、 $3/4$ 、 $2/4$ と計算され、その結果、入力単語列（「私」、「学校」、「行く」）と用例「これは木でできている」との類似度は  $8/4$ （ $3/4 + 3/4 + 2/4$ ）と求められる。

【0062】以上のようにして、用例データベース8に登録されているすべての用例について、複数の文候補（入力単語列）それぞれに対する類似度が計算される。

【0063】その後、用例検索部5では、類似度が最も高い用例と入力単語列とが選択される（以下、選択された用例または入力単語列を、それぞれ選択用例または選択単語列という）。

【0064】そして、用例検索部5は、選択用例を構成する単語のうち、選択単語列を構成する単語に対応するものを、その選択単語列を構成する単語に、それぞれ置き換え、文音声認識結果として出力する。

【0065】即ち、例えば、いま、説明を簡単にするために、用例として「彼は会社を訪問する」および「これは木でできている」が用例データベース8に記憶されており、入力単語列として、（「私」、「学校」、「行く」）が得られたとすると、上述したように、入力単語列（「私」、「学校」、「行く」）と、用例「彼は会社を訪問する」または「これは木でできている」それぞれの類似度は、 $3/4$ または $8/4$ となる。本実施例では、類似度が高い場合というのは、類似度の値が小さい場合であり（これは、図2において、シソーラスを構成する最下層の概念の階層から、第1階層、第2階層、・・・としたためで、これとは逆に、シソーラスを構成する最上層の概念の階層から、第1階層、第2階層、・・・とすれば、類似度が高い場合というのは、類似度の値が大きい場合となる）、従って、用例選択部5では、入力単語列（「私」、「学校」、「行く」）と用例「彼は会社を訪問する」とが選択される。

【0066】そして、用例検索部5では、選択用例「彼は会社を訪問する」を構成する単語のうち、選択単語列（「私」、「学校」、「行く」）を構成する単語に対応するもの、即ち「私」、「学校」、「行く」に対応する「彼」、「会社」、「訪問する」が、その選択単語列を構成する単語「私」、「学校」、「行く」に、それぞれ置き換えられ、さらに必要に応じて助詞の変換がなされ、その結果得られる「私は学校に行く」が文音声認識

結果として出力される。

【0067】この音声認識結果は、翻訳部11に供給され、翻訳部11では、用例検索部5からの音声認識結果「私は学校に行く」が、対訳辞書12を参照して、例えば英文「I go to school」に英訳される。そして、この英訳文「I go to school」に対応する合成音が生産されて出力される。

【0068】なお、以上においては、音声が日本語で入力された場合について説明したが、英語で、例えば音声「I go to school」が入力された場合についても、音声認識装置10または翻訳装置20において、上述の場合と同様にして音声認識または翻訳がなされ、その結果得られる「私は学校に行く」が合成音で出力される。

【0069】ところで、用例検索部5では、上述のような類似度の計算にあたって、いままでに選択された用例による文脈に基づく重みが付されるようになされている。即ち、用例データベース8には、用例とともに、例えば今回選択された用例から考えて次の音声認識処理に用いられる可能性のある用例を予測するための予測情報が記憶されており、用例検索部5では、その予測情報に基づいて、類似度に重みが付けられるようになされている。

【0070】具体的には、用例データベース8には、予測情報として、用例を次の音声認識処理に用いられる（次の音声認識処理で選択される）可能性のある用例と対応付けたものが記憶されており、用例検索部5は、前回の音声認識処理で用いられた用例と対応付けられている用例に対する類似度が高くなるように、または前回の音声認識処理で用いられた用例と対応付けられていない用例に対する類似度が低くなるように、重みを付けるようになされている。

【0071】ここで、用例データベース8には、用例を、次の音声認識処理に用いられる可能性のある用例と対応付けたものが、例えば図3に示すように表形式で記述されて記憶されている。この表は、その左欄に記述された用例が前回用いられた場合に、今回用いられる可能性の高い用例を右欄に記述したもので、左欄から右欄の方向に用例を見た場合に、その2つの用例による文脈が、人の発話または対話において現れ得るようなものとされている。従って、図3の表の左欄から右欄の方向は、発話または対話の流れを表しているといえることができる。以下、適宜、このような発話または対話の流れを表現する、いわば脚本というべき表の左欄と右欄との組み合わせの記述を、スクリプトという。

【0072】図3の実施例では、用例「チェックインをお願いします。」に続いて用いられる可能性の高い用例として、例えば「お客様のお名前をどうぞ。」、「ご予約はされていますか。」、および「この用紙にご記入をお願いします。」が登録されている。また、用例「部屋はありますか。」に続いて用いられる可能性の高い用例

として、例えば「ございます。」、「ございません。」、および「満室です。」が登録されている。

【0073】例えば、いま、図3に示した表が用例データベース8に登録されている場合において、米国のホテルのフロントで、日本語で「チェックインをお願いします。」と発話し、音声認識装置10において、入力音声「チェックインをお願いします。」が、用例「チェックインをお願いします」を用いて音声認識されたとする。この場合、音声認識装置10からは、音声認識結果として、「チェックインをお願いします。」が出力され、これが、翻訳部11で英訳され、例えば「Checkin, please.」が合成音で出力される。

【0074】そして、この合成音を聞いたホテルのフロントの者が、英語で「お客様のお名前をどうぞ。」と発話したとする（例えば、「May I have your name, please.」と発話したとする）。この入力音声は、音声認識装置10において、上述したように類似度を最も高くする用例が選択されて音声認識されるが、この場合、類似度には、スクリプトを参照して、前回の音声認識処理で用いられた用例「チェックインをお願いします。」と対応付けられている用例（図3に示した実施例においては、「お客様のお名前をどうぞ。」、「ご予約はされていますか。」、および「この用紙にご記入をお願いします。」）に対する類似度が高くなるように、重みが付けられる。

【0075】即ち、この場合、音声認識装置10の用例検索部5では、図4のフローチャートに示すように、まず最初に、ステップS1において、前回の音声認識処理で用いられた用例「チェックインをお願いします。」と対応付けられている用例（スクリプト）「お客様のお名前をどうぞ。」、「ご予約はされていますか。」、および「この用紙にご記入をお願いします。」が選択され、ステップS2に進み、上述したように、シソーラスを用いて、認識部4が出力する複数の文候補それぞれと、用例データベース8に記憶されている複数の用例それぞれの類似度が計算される。

【0076】そして、ステップS3に進み、ステップS1で選択された用例（スクリプト）に対応する類似度が高くされ、ステップS4に進む。ステップS4では、最も類似度を高くする用例および文候補が選択され、以下、上述したようにして最終的な音声認識結果が求められる。

【0077】従って、ステップS2の段階で、用例「チェックインをお願いします。」に対する類似度が最も高くなくても、ステップS3の処理により、その類似度が高くされ、その結果、用例「チェックインをお願いします。」に対する類似度を最も高くすることができる。

【0078】以上のように、単語のシソーラス上での意味上の類似性のみだけでなく、前回用いられた用例による文脈に基づいて、対話（あるいは発話）の流れを考慮

して、類似度を求めるようにしたので、従来の場合より、音声認識処理を行うのに適切な用例を得ることができ、その結果、音声認識処理の精度、即ち音声の認識率を向上させることができる。

【0079】なお、図3において、表の右欄に記述した用例は、ホテルのフロントの者が発話すると予測されるものであるから、実際には、英語で記述されている（例えば、「お客様のお名前をどうぞ。」は、「May I have your name, please.」と記述されている）。

【0080】また、上述の場合には、前回の音声認識処理で用いられた用例と対応付けられている用例に対する類似度が高くなるように、類似度に重みを付けるようにしたが、この他、類似度には、前回の音声認識処理で用いられた用例と対応付けられていない用例に対する類似度が低くなるように、重みを付けるようにすることも可能である。

【0081】さらに、上述の場合においては、前回の音声認識処理で用いられた用例と対応付けられている用例に対する類似度が高くなるように、類似度に重みを付けるようにしたが、この他、さらにそれ以前の処理で用いられた用例に基づいて、類似度に重みを付けるようにすることも可能である。即ち、例えば宿泊客が、既に名乗っており、そのような音声の認識のための用例を既に用いている場合には、相手の名前を尋ねるときに発する音声を認識するための用例（例えば、図3に示した「お客様のお名前をどうぞ。」など）に対する類似度を低くする重みを付けるようにすることが可能である。

【0082】また、用例データベース8には、図3に示した表（スクリプト）を、例えばホテルや、レストラン、空港などの所定の場面ごとに用いるべきものに分類して記憶させておき、用例検索部5には、図1の音声翻訳装置が用いられる場面に応じた表（スクリプト）を用いて、処理を行わせるようにすることができる。この場合、音声認識処理に用いるスクリプトが限定されるので、さらなる音声認識性能の改善および計算量の削減を図ることができる。なお、音声翻訳装置がどのような場面で利用されているかは、図示せぬ操作部を操作することにより、装置に認識させるようにすることが可能である。また、発話の内容から、音声翻訳装置がどのような場面で利用されているかを、装置に認識させるようにすることもできる。即ち、例えば音声「部屋はありますか。」などが入力された場合には、ホテルでの発話または対話の流れを反映した表を用いて、また、例えば音声「メニューを見せて下さい。」などが入力された場合には、レストランでの発話または対話の流れを反映した表を用いて、用例検索部5に処理を行わせるようにすることができる。

【0083】さらに、用例データベース8には、次の処理に用いられる可能性のある用例を学習により求めたものを記憶させておくようにすることができる。この場

合、学習の際に、大量の学習用データを用いて、次の処理で、各用例が用いられる確率を求めておくようにし、その確率に応じて、類似度を高くするまたは低くする重みを調整することで、より処理に適した用例の選択を行うことが可能となる。

【0084】また、翻訳部 11 における翻訳の方法については特に言及しなかったが、翻訳部 11 では、例えば上述した特開平 3-276367 号公報に開示されている翻訳方法により翻訳を行うようにすることが可能である。この場合、用例データベース 8 は、用例検索部 5 と翻訳部 11 とで共用することが可能である。さらに、翻訳部 11 では、用例検索部 5 における場合と同様に、類似度の計算にあたって、いままでに選択された用例による文脈に基づく重みを付すようにすることも可能である。この場合、翻訳の精度を向上させることができる。なお、翻訳部 11 における翻訳の方法は、特開平 3-276367 号公報に開示されているものに限定されるものではない。

【0085】次に、図 5 は、本発明を適用した音声翻訳装置の他の実施例の構成を示している。なお、図中、図 1 における場合と対応する部分については、同一の符号を付してある。即ち、この音声翻訳装置は、認識部 4 または用例検索部 5 に代えて、それぞれ認識部 14 または用例検索部 15 が設けられている他は、図 1 の音声翻訳装置と同様に構成されている。

【0086】認識部 14 は、例えば図 1 の認識部 4 における場合と同様にして音声認識を行い、複数の文候補を、用例検索部 15 に出力するようになされている。さらに、認識部 14 は、複数の文候補それぞれに加えて、その確からしさを示す尤度（例えば、HMM により音声認識を行う場合には、上述した生起確率）も、用例検索部 5 に出力するようになされている。

【0087】用例検索部 15 は、やはり図 1 の用例検索部 5 と同様に、認識部 14 から供給される複数の文候補それぞれを入力単語列に変換し、用例データベース 8 に記憶されている用例それぞれとの類似度を計算して、その類似度を最も高くする用例を選択するようになされている。そして、その選択した用例に基づいて、音声入力部 1 に入力された音声の認識結果（文）を決定するようになされている。但し、用例検索部 15 における類似度

$$S_g = \alpha S_1 + (1 - \alpha) S_s$$

但し、 $\alpha$  は、 $0 < \alpha < 1$  の範囲のいずれかの実数値をとる。

【0094】式 (1) にしたがって、総合類似度を算出した後は、ステップ S14 に進み、総合類似度を最も高くする用例および文候補が選択され、以下、図 1 における場合と同様の処理が行われることにより、翻訳結果が出力される。

【0095】なお、式 (1) による総合的類似度の計算は、複数の文候補それぞれについて、言語的類似度を最

の計算は、認識部 14 から供給される複数の文候補それぞれの尤度を考慮して行われるようになされている。

【0088】以上のように構成される音声翻訳装置では、音声が入力されると、音声入力部 1、AD 変換部 2、分析部 3、および認識部 14 において、図 1 の音声翻訳装置と同様の処理が行われ、これにより認識部 14 からは、複数の文候補が、用例検索部 15 に出力される。さらに、認識部 14 からは、複数の文候補それぞれの尤度も用例検索部 15 に出力される。

【0089】用例検索部 15 では、図 6 のフローチャートに示すように、まず最初に、ステップ S11 において、認識部 14 からの複数の文候補（文認識候補）と、それぞれの尤度が受信され、入力単語列に変換される。そして、ステップ S12 に進み、複数の文候補それぞれと、用例データベース 8 に記憶されている用例それぞれとの類似度が、シソーラス、さらに必要に応じてスクリプトを用いて計算される（スクリプトを用いることは必須ではないが、用いた方が、類似度の精度が向上させることができる）。

【0090】ここで、ステップ S12 で求められる類似度は、文法的、意味的な正しさを表しており、即ち、いわば言語的な正しさを表しており、以下、適宜、言語的類似度という。また、これに対し、認識部 14 から出力される尤度は、音響的な処理の結果得られた複数の認識結果候補それぞれの正しさを表しており、以下、適宜、音響的類似度という。

【0091】その後、ステップ S13 に進み、複数の認識結果候補それぞれと、用例データベース 8 に記憶されている複数の用例それぞれとの総合的な類似度である総合類似度が、言語的類似度および音響的類似度に基づいて計算される。

【0092】ここで、総合類似度は、例えば言語的類似度および音響的類似度を用いて表現される多項式にしたがって計算される。即ち、総合類似度は、言語的類似度と音響的類似度とを、所定の割合で加算することにより計算される。具体的には、例えば、いま総合類似度、言語的類似度、または音響的類似度を、それぞれ  $S_g$ 、 $S_1$ 、または  $S_s$  とするとき、総合類似度  $S_{gi}$  は、次式にしたがって計算される。

【0093】

$$\dots (1)$$

も高くする用例についてだけ行えば良い。また、総合的類似度の計算は、言語的類似度および音響的類似度を正規化してから行うのが好ましい。

【0096】以上のように、最終的な音声認識結果を得るのに用いる用例を選択する際に、言語的類似度だけでなく、認識部 4 において得られる音響的類似度（尤度）を考慮するようにしたので、言語的類似度だけを考慮する場合に比較して、より適切な用例を選択することが可能となり、その結果、音声の認識率を向上させることが

できる。

【0097】即ち、逆にいえば、最終的な音声認識結果を得るのに、音響的類似度だけでなく、言語的類似度を考慮するようにしたので、何らかの原因で（例えば、雑音などにより）、本来正しい文候補の音響的類似度が低くなっても、言語的類似度を考慮することで、その総合類似度を高くすることができ、本来正しい文候補が誤ったものであるとされることを防止することが可能となる。

【0098】なお、式（1）における言語的類似度と音響的類似度とを加算するときの割合である $\alpha$ は、動的に変化させることができる。即ち、音声翻訳装置が使用される環境における雑音レベルが高い場合には、音響的類似度の信頼性が低下すると考えられる。そこで、周囲の雑音レベルを検出するようにし、それが高い場合には、 $\alpha$ の値を大きくすることで、総合類似度に対し、音響的類似度が寄与する割合を低下させるとともに、言語的類似度が寄与する割合を増加させるようにする。また、雑音レベルが低い場合には、 $\alpha$ の値を元に戻すようにする。このようにすることで、環境の変化に対する装置の耐性を向上させることができる。

【0099】以上のように、類似度（総合類似度）の計算にあたって、シソーラスに基づく類似度に加えて、いままで用いられた用例による文脈や、音響的類似度を考慮するようにしたので、用例の選択の精度、ひいては音声処理（自然言語処理）の精度を向上させることができる。

【0100】以上、本発明を音声翻訳装置に適用した場合について説明したが、本発明は、その他、例えば音声認識装置や機械翻訳装置などの、自然言語処理を必要とするあらゆる装置に適用可能である。

【0101】なお、本実施例においては、日本語を英語に翻訳するようにしたが、日本語以外の言語を翻訳したり、また、英語以外の言語に翻訳するようにすることが可能である。

【0102】また、本実施例では、認識部4（認識部14についても同様）において、言語モデルによる制約の下、文候補を得て、用例検索部5（用例検索部15についても同様）において、それを単語列に分離するようにしたが、この他、認識部4では、例えば言語モデルによる制約なしで連続音声認識を行い、用例検索部5では、その結果得られる単語を組み合わせる単語列を構成するようにすることも可能である。

【0103】さらに、本実施例では、入力単語列に助詞を含ませないようにしたが、用例検索部5には、助詞を含ませた入力単語列を処理させるようにすることも可能である。

【0104】また、本実施例においては、翻訳結果を合成音で出力するようにしたが、この他、翻訳結果は、表

示するようにすることも可能である。

【0105】

【発明の効果】本発明の自然言語処理装置および自然言語処理方法によれば、用例記憶手段に記憶されている用例それぞれと、入力文との類似度を計算し、その類似度を最も高くする用例を選択するときに、いままでに選択した用例による文脈に基づいて重みを付けて、類似度が計算される。従って、類似度の精度が向上するので、より処理に適した用例を選択することが可能となる。

【0106】本発明の音声認識装置および音声認識方法によれば、複数の認識結果候補と、それぞれの尤度とが求められ、複数の認識結果候補それぞれと、用例記憶手段に記憶されている用例それぞれとの言語的な類似度である言語的類似度が計算される。さらに、複数の認識結果候補それぞれと、用例記憶手段に記憶されている用例それぞれとの総合的な類似度である総合類似度が、言語的類似度および尤度に基づいて計算される。そして、総合類似度を最も高くする用例が選択され、その用例に基づいて、音声の認識結果が求められる。従って、音声の認識率を向上させることができる。

【図面の簡単な説明】

【図1】本発明を適用した音声翻訳装置の第1実施例の構成を示すブロック図である。

【図2】シソーラスを示す図である。

【図3】用例データベース8に記憶されている表を示す図である。

【図4】用例検索部5の動作を説明するフローチャートである。

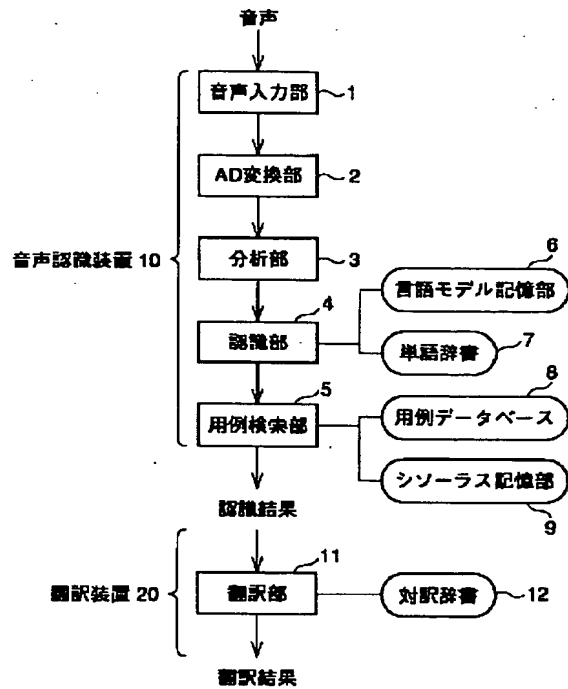
【図5】本発明を適用した音声翻訳装置の第2実施例の構成を示すブロック図である。

【図6】用例検索部15の動作を説明するフローチャートである。

【符号の説明】

- 1 音声入力部
- 2 AD変換部
- 3 分析部
- 4 認識部
- 5 用例検索部
- 6 言語モデル記憶部
- 7 単語辞書
- 8 用例データベース
- 9 シソーラス記憶部
- 10 音声認識装置
- 11 翻訳部
- 12 対訳辞書
- 14 認識部
- 15 用例検索部
- 20 翻訳装置

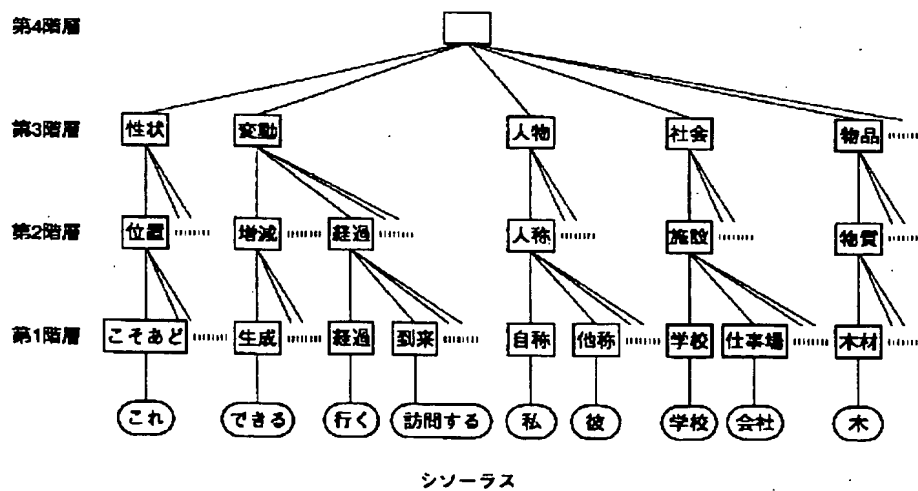
【図1】



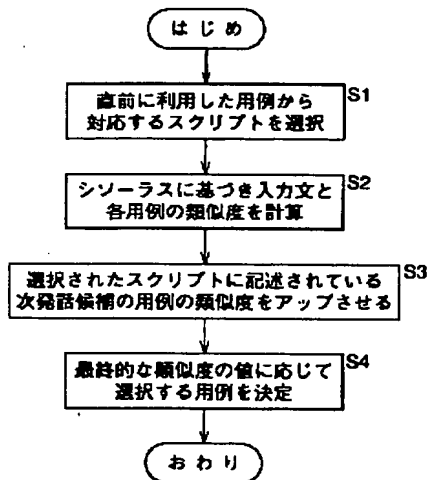
【図3】

直前の用例	次に利用される可能性の高い用例
チェックインをお願いします。	お客様のお名前をどうぞ。 ご予約はされていますか。 この用紙にご記入をお願いします。
部屋はありますか。	ございます。 ございません。 満室です。
⋮	⋮

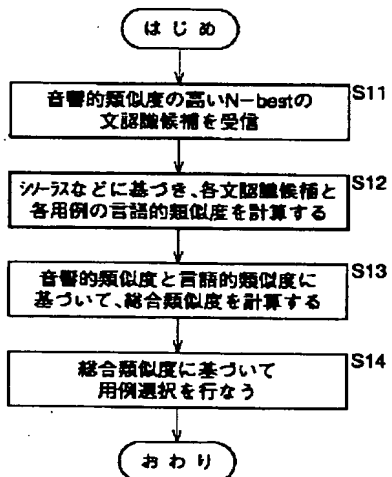
【図2】



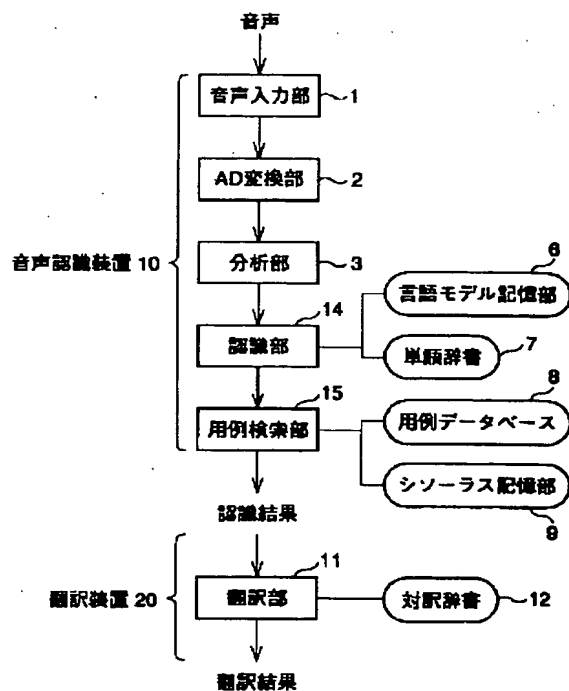
【図 4】



【図 6】



【図 5】



フロントページの続き

(72) 発明者 加賀美 徹也  
東京都品川区北品川 6 丁目 7 番 35 号 ソニ  
ー株式会社内

(72) 発明者 石井 和夫  
東京都品川区北品川 6 丁目 7 番 35 号 ソニ  
ー株式会社内

(72) 発明者 田中 幸  
東京都品川区北品川 6 丁目 7 番 35 号 ソニ  
ー株式会社内

(72) 発明者 加藤 靖彦  
東京都品川区北品川 6 丁目 7 番 35 号 ソニ  
ー株式会社内

(72) 発明者 角田 弘史  
東京都品川区北品川 6 丁目 7 番 35 号 ソニ  
ー株式会社内

(72) 発明者 小川 浩明  
東京都品川区北品川 6 丁目 7 番 35 号 ソニ  
ー株式会社内